

Studi & Ricerche

STUDIES & RESEARCH

Michele Lanzetta, Federica Fanti, Giovanni Tantussi*

- **Aspetti critici** della marcatura di materiali lapidei con **caratteri alfanumerici**
Critical aspects of marking stone materials with alphanumerical characters
- **La visione artificiale** permette l'**automazione** del riconoscimento dei codici
Artificial vision makes it possible to automate code recognition
- **Classificazione** di approcci per l'OCR proposti in letteratura
Classification of the OCR approaches offered in literature

Automazione e controllo visivo della marcatura di prodotti (Parte Seconda)

Automating and visually inspecting product marking (Part Two)

Da cosa è costituito un carattere?
Questa è la domanda a cui i ricercatori hanno cercato di rispondere con lo studio e lo sviluppo della **segmentazione** (*segmentation*), importante fase nello studio della leggibilità [14] [15] [16]. La segmentazione è una delle fasi nel

riconoscimento dei caratteri ed è una operazione di decomposizione dell'immagine in una sequenza di simboli individuali (figura 6). Questo è uno dei tipici processi di decisione che avvengono nei sistemi di riconoscimento dei caratteri (OCR). Il sistema deve decidere se il simbolo isolato è un'unità riconoscibile, in caso contrario verrà appurato l'errore e avverrà una nuova ricerca. Il modo di segmentare i caratteri non è una decisione locale, indipendente dalle decisioni precedenti e seguenti. Avere un buon database di caratteri è necessario ma non sufficiente per un buon riconoscimento. Anche una serie di "pattern" (ad es. dittonghi) soddisfacenti può essere giudicata scorretta se le richieste contestuali del sistema di

What does a character consist of?
This is the question researchers have tried to answer by studying and developing **segmentation**, an important term in studying legibility [14] [15] [16]. Segmentation is one of the stages of character recognition and is the operation of decomposing the image into a sequence of individual symbols (figure 6). This is one of the typical decisional processes taking place in optical character recognition systems (OCR). The system has to decide if the isolated symbol is a recognizable unit and, if not, will check the error and start a new search.

The way in which characters are segmented is not a local decision, independent of prior and later decisions. Having a good database of characters is necessary but not enough for good recognition. Even a set of satisfactory "patterns" (diphthongs, for example) may be judged incorrect if the output system's contextual

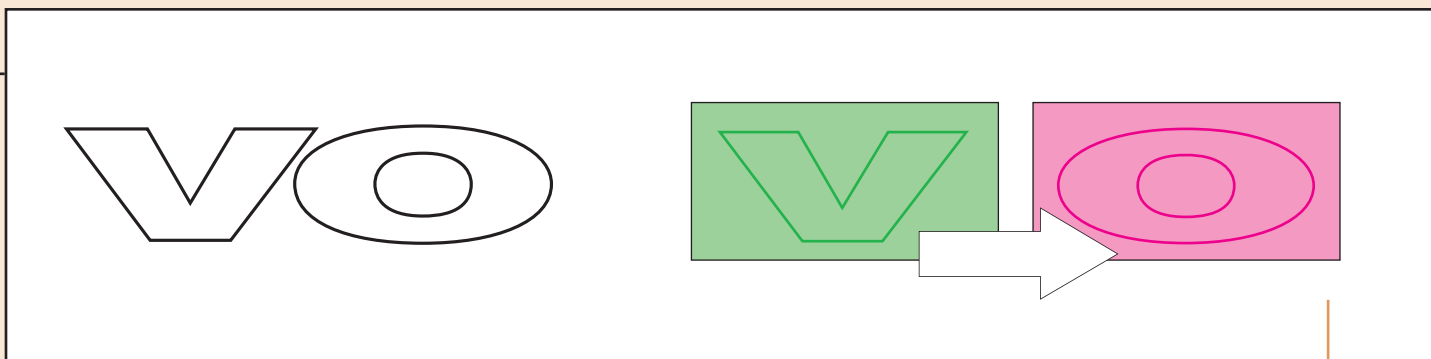


Fig. 6. In caso di crenatura, cioè di riduzione dello spazio tra due caratteri, come nella coppia a sinistra, non è possibile una segmentazione attraverso rettangoli come in assenza (coppia di caratteri a destra), ma sono necessari poligoni per isolare l'area che contiene un singolo carattere.
In the case of kerning, that is reducing the space between two characters, as in the pair on the left, segmentation cannot be done through rectangles as in absence (pair on the right); polygons are needed to isolate the area containing a single character.

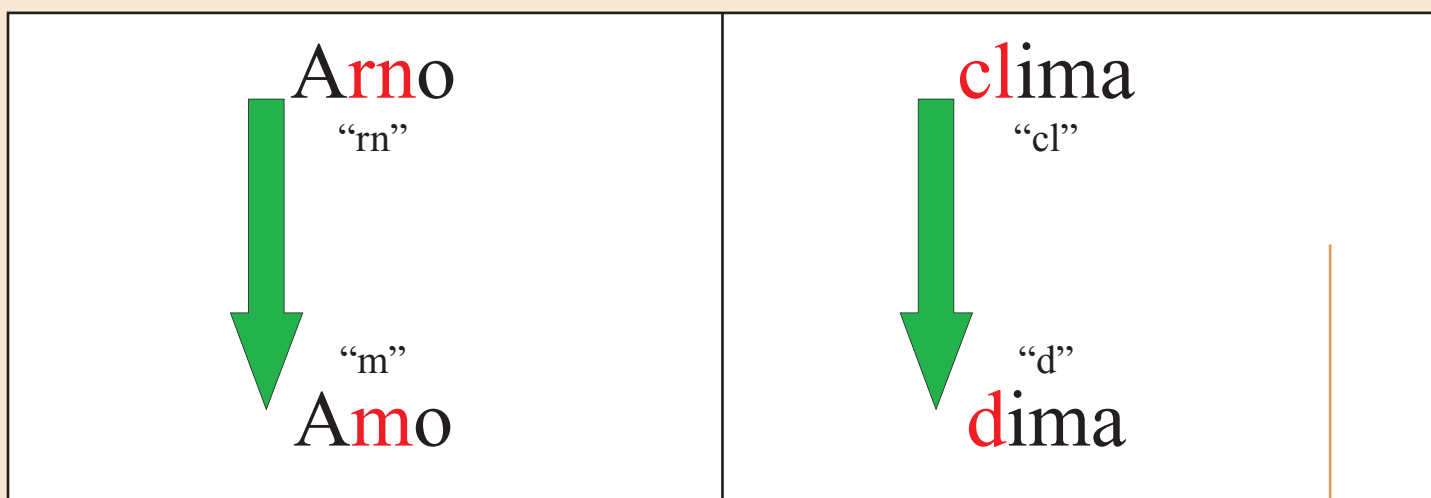


Fig. 7. Esempi di errori di leggibilità che possono propagarsi nel sistema.
Examples of legibility errors that can propagate in the system.

uscita non sono soddisfatte. Per esempio [17], la sequenza di lettere "cl" può essere letta dal sistema come una "d" (figura 7). Anche se esistono altri strumenti *a posteriori* che scartano la possibilità di un risultato valido sfruttando la sfera semantica, non è comunque detto che questi sistemi non siano affetti da errore, infatti alcuni errori commessi al primo passo possono amplificarsi se si formano parole contenute comunque nel database.

Word segmentation

L'analisi strutturale del layout è effettuata per ottenere la segmentazione fisica in gruppi costituenti parole isolate, linee di testo e blocchi di linee come paragrafi separati. Esistono due modi di affrontare questo tipo di analisi, ovvero la *top-down analysis* (figure 8) in cui una pagina viene suddivisa dal componente più grande a quello più piccolo, per es. da linee di testo a caratteri, e la *bottom-up analysis* in cui i componenti connessi sono fusi in caratteri, questi in parole e di seguito in linee di testo. Questi due modelli di analisi hanno dato vita ad una serie di modelli ibridi che sono più robusti

requirements have not been met. For example [17], the "cl" letter sequence could be read by the system as a "d" (figure 7). And while other *a posteriori* systems exist that discard the possibility of a valid result exploiting the field of semantics, it is not necessarily true that they are error-free; in fact some first-step errors can even amplify if they form words in any case contained in the database.

Word segmentation

A structural analysis of the layout is done to get physical segmentation into groups making up isolated words, rows of text and blocks of rows as separate paragraphs. There are two ways of dealing with this type of analysis: *top-down analysis* (figure 8), in which the page is divided up by the largest to the smallest component, e.g. from rows of text to characters; and *bottom-up analysis* in which the connected components are merged into characters, characters into words and words into rows of text. These two analysis models have led to a series of hybrid

Studi & Ricerche

STUDIES & RESEARCH

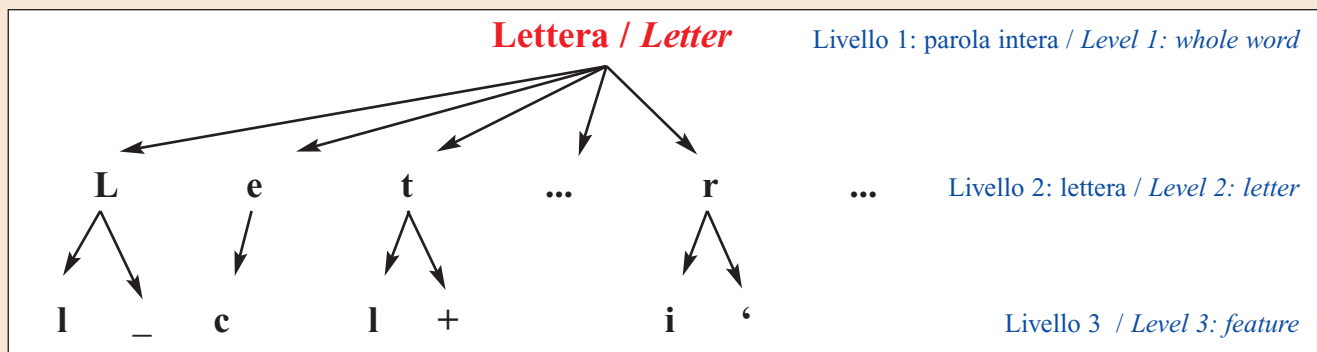


Fig. 8. Diagramma dei livelli di segmentazione.
Diagram of segmentation levels.

eusano un numero maggiore di informazioni sul testo. È intuitivo pensare che la distanza fra i caratteri sia minore di quella fra le parole. Questa è l'informazione che viene usata per fondere i caratteri in parole. La segmentazione delle parole è molto importante per la sintesi di un discorso. Un buon modo, usato per dividere una parola è **inter-word distance IWD** [18] definito come:

$$IWD > std(ICD) + mean(ICD)$$

Dove con ICD si intende la distanza fra i caratteri (*inter-character distance*) e con $std(ICD)$ e $mean(ICD)$ rispettivamente la sua deviazione standard e la sua media. La loro quantificazione avviene dopo aver ridefinito la segmentazione, per avere una deviazione standard e la media aggiornate in modo iterativo, al fine di diminuire gli errori.

Esistono metodi sofisticati che sono stati sviluppati negli ultimi anni in cui il riconoscimento è basato sulla segmentazione. In Casey e Lecolinet [17] la distinzione fra i vari metodi di segmentazione presenti in letteratura avviene per il criterio secondo cui la segmentazione e la classificazione interagiscono con l'intero sistema. Alla base di tutti i metodi comunque c'è il fatto che un riconoscimento inaccettabile viene riesaminato, modificato e segmentato nuovamente. Una profonda interazione fra i due aspetti del riconoscimento si verifica quando la classificazione avviene attraverso una scelta fra un set di possibilità. In questo tipo di approccio la segmentazione e la classificazione sono integrate.

La classificazione proposta in [17] può essere interpretata come un diagramma cartesiano (figura 9) in cui gli assi rappresentano i tre metodi considerati "puri", mentre nello spazio si trovano i metodi che nascono dalla fusione di caratteristiche dei tre mescolate in modo differente. I tre metodi detti puri schematizzati in figura 9 sono così caratterizzati:

1. L'approccio classico, in cui i segmenti identificati sono basati sulle proprietà dei caratteri (*character-like*).

*models that are sturdier and use more data about the text. It is intuitive to think that the distance between characters is smaller than the distance between words. This is the information used to merge characters into words. Word segmentation is very important in synthesizing a discourse. A good way of dividing up a word is **inter-word distance IWD** [18] defined as:*

$$IWD > std(ICD) + mean(ICD)$$

Where ICD means inter-character distance and $std(ICD)$ and $mean(ICD)$ respectively stand for standard deviation and its mean. They are quantified after having redefined segmentation in order to get standard deviation and the mean updated in an iterative way so as to lessen error.

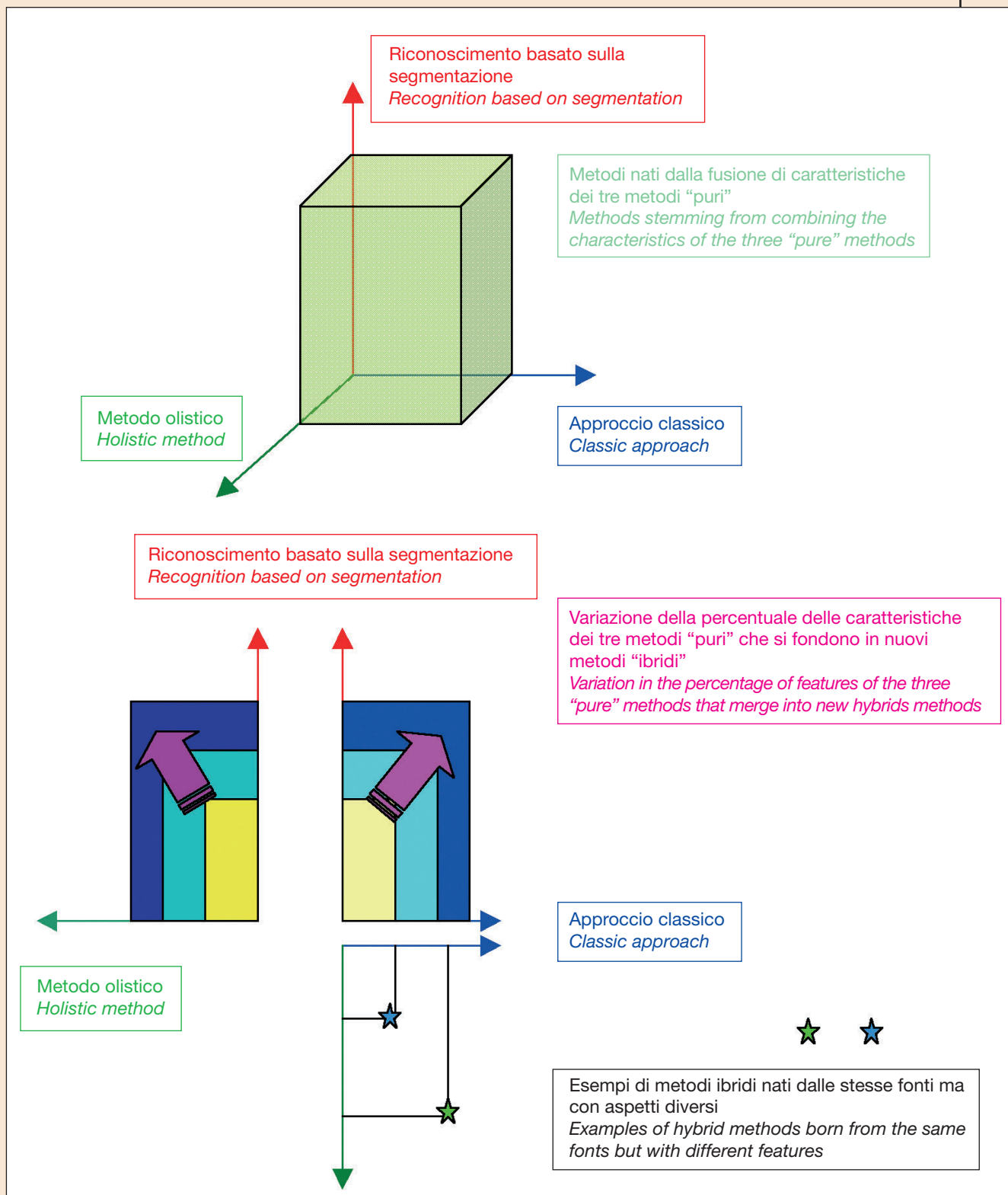
Sophisticated methods have recently been developed basing recognition on segmentation. In Casey and Lecolinet [17] a distinction between the various segmentation methods described in literature is made according to the criterion of how segmentation and classification interact with the entire system. In any case, at the basis of all the methods is the fact that unacceptable recognition is re-examined, modified and newly segmented. Profound interaction between the two aspects of recognition occurs when classification is done by means of a choice between a set of possibilities. In this type of approach segmentation and classification are integrated.

The classification proposed in [17] can be interpreted as a Cartesian diagram (figure 9) in which the axes represent the three methods considered "pure", while in the space are found the methods stemming from combining the characteristics of the three in different ways.

The three so-called pure methods schematized in figure 9 are characterized as:

1. *The classic approach, in which identified segments are based on character-like features. This process of dividing the image into significant components is called "dissection".*

Fig. 9. Diagramma Cartesiano che mostra la teoria formulata da Casey e Lecolinet in [17].
Cartesian diagram showing the theory formulated by Casey and Lecolinet in [17].



Studi & Ricerche

STUDIES & RESEARCH

Questo processo di suddivisione dell'immagine in componenti significative prende il nome di "dissezione".

2. Il riconoscimento basato sulla segmentazione, nel quale il sistema ricerca nell'immagine componenti in base a classi diverse presenti nel suo alfabeto.
3. I metodi olistici, nei quali il sistema riconosce le parole come un intero, eliminando il bisogno di suddividere in caratteri.

Il primo metodo sceglie, come criterio per una buona segmentazione, un modo che va d'accordo con le proprietà generali dei segmenti ottenuti e con i caratteri considerati validi.

Nel secondo, il criterio è il riconoscimento delle semplici unità da cui è costituito il carattere stesso, includendo però la correttezza sintattica o semantica del risultato totale.

Il metodo olistico verte in pratica sull'approccio classico di riconoscimento delle lettere dell'alfabeto.

Il metodo ibrido di *oversegmenting* prevede due strategie differenti di segmentazione: la prima differenza riguarda la scelta dei punti di segmentazione da un'analisi generale delle immagini; la seconda è all'estremo opposto, qui non viene compiuta nessuna dissezione, gli algoritmi di classificazione semplicemente danno una forma al modello senza considerare il contenuto dell'immagine.

In questo caso viene applicato all'immagine un algoritmo di dissezione chiamato "oversegmenting", cioè l'immagine viene tagliata in un numero di parti sufficienti in modo tale che i legami corretti siano compresi tra tutti i tagli eseguiti. Una volta che la robustezza dell'algoritmo è stata assicurata, la segmentazione ottimale è definita da un determinato set di tagli eseguiti. Ogni sottoinsieme (*subset*) implica un'ipotesi di segmentazione mentre la classificazione è portata a sostenere ipotesi differenti e a scegliere la segmentazione migliore. Ogni possibile segmentazione è testata separatamente dalla classificazione e la suddivisione che dà il più alto riconoscimento viene accettata. Ogni ipotesi di segmentazione è generata in un singolo passo. Quando il numero dei caratteri nell'immagine da sezionare non è conosciuto a priori, le varie ipotesi sono di solito generate in due passi. Nel primo viene determinato un set di tagli possibili e l'immagine in ingresso è divisa in componenti elementari da separare lungo ogni percorso (prima stabilito). Nella seconda parte le ipotesi di separazione sono generate dalle combinazioni formate con le componenti elementari. Tutte le combinazioni incontrano costrizioni di accettabilità e sono il risultato di una **classificazione**.

Automazione del controllo della marcatura di prodotti industriali

Nei paragrafi precedenti sono state classificate le principali fasi del processo di OCR e sono state indicate le difficoltà presenti quando non si ha a che fare con la semplice carta stampata, pertanto le classi di problematiche possibili dipendono da:

1. contrasto del tratto che compone un carattere
2. leggibilità.

Le problematiche di leggibilità possono essere affrontate

2. *Recognition based on segmentation, in which the system searches for image components on the basis of different classes in its alphabet.*

3. *The holistic methods, in which the system recognizes words in their entirety, eliminating the need to divide them into characters.*

As a criterion for good segmentation, the first method chooses a way in agreement with the general features of the segments obtained and with the characters considered valid.

In the second the criterion is recognizing simple units from which the character is composed, including, however, the syntactical or semantic correctness of the total result.

The holistic method hinges in practice on the classic approach of recognizing letters of the alphabet.

The hybrid method of oversegmenting envisions two different segmentation strategies: the first difference concerns choosing segmentation points through general image analysis; the second is the exact opposite, with no dissection but with the classification algorithms simply giving form to the model without considering image content.

In this case, applied to the image is a dissection algorithm called "oversegmenting": in other words, the image is cut into enough parts so that the correct ties are included between all the cuts made.

*Once the sturdiness of the algorithm has been assured, optimal segmentation is defined by a given set of cuts. Each subset implies a segmentation hypothesis while classification is led to uphold different hypotheses and choose the best segmentation. Each possible segmentation is tested separately by classification and the subdivision getting the highest recognition is the one accepted. When the number of characters in the image to be divided is not known a priori the various hypotheses are usually generated in two steps. The first determines a set of possible cuts and the input image is divided into elementary components to be separated along each (preset) route. In the second the separation hypotheses are generated by the combinations formed with the elementary components. All the combinations run into acceptability constrictions and are the result of a **classification**.*

Automating inspection of industrial product marking

The previous paragraphs classified the main stages of the OCR process and showed the difficulties that can arise when not dealing with simple printed paper, and so the classes of possible problems depend on:

1. *contrast of the feature composing a character*
2. *legibility.*

Legibility problems can be dealt with using methods similar to those described, from many standpoints well established from the scientific and industrial standpoints, given the presence of widely used commercial products such as OCR software for PC

Studi & Ricerche

STUDIES & RESEARCH

con metodi simili a quelli descritti e per molti aspetti sono consolidate sia dal punto di vista scientifico sia da quello industriale, data la presenza di prodotti commerciali di grande diffusione, come i software di OCR per PC, e considerata la presenza di apposite funzioni OCR nei comuni sistemi di visione artificiale.

Viceversa, il problema di individuare il tratto che costituisce il carattere su prodotti marcati con diverse tecnologie ha un crescente interesse e richiede una definizione più accurata.

Lo scopo finale del presente articolo è quello di sottolineare che la forma dei caratteri stampati è definita anche da norme internazionali, quali [19] [20] [21], che permettono una agevole definizione di algoritmi per il riconoscimento, mentre il controllo di un processo di marcatura o lo sviluppo di una nuova tecnologia può essere semplicemente affrontato esaminando il tratto che viene generato nella marcatura del carattere stesso. Un esempio di misura del contrasto di un carattere rispetto ad uno sfondo (con numerosi disturbi) può essere dato dal rapporto tra il segnale del carattere e quello dello sfondo, basato sulla estensione geometrica del tratto nell'area inquadrata e del livello di intensità del segnale. Tale esempio con molti altri può essere definito matematicamente e utilizzato come parametro di valutazione delle prestazioni di algoritmi di riconoscimento caratteri in ambito industriale. Relativamente ad una immagine monocromatica, si può ad esempio considerare il rapporto tra l'intensità media dei pixel appartenenti al carattere e quella media dello sfondo. Maggiore è la variabilità o il rumore presente sullo sfondo, minore risulterà il contrasto e quindi la leggibilità.

Le due problematiche citate risultano fortemente integrate, in quanto le prestazioni di un algoritmo di OCR per qualunque fase del processo (schematizzate in figura 1, nella prima parte di questo articolo) sono determinate dalla tecnologia di marcatura e quindi dal tipo di immagine acquisita.

Conclusioni

È stata presentata una panoramica dei principali filoni di ricerca trattati dagli sviluppatori di sistemi automatici di marcatura e controllo.

Il problema del riconoscimento caratteri (OCR) è ampiamente trattato nella ricerca informatica, ma dal punto di vista industriale le problematiche richiedono ancora un inquadramento sistematico in modo da favorire la realizzazione di software commerciali in grado di affrontare la fase di preprocessing in modo affidabile per una vasta gamma di tecnologie di marcatura su materiali diversi dalla carta.

Nell'ultima parte di questo articolo è stato introdotto il concetto di contrasto del tratto. In una marcatura di prodotti il problema del contrasto necessita una apposita trattazione ai fini sia del **controllo della qualità** delle marcature eseguite, sia della definizione dei requisiti per un **sistema automatico**. Ricerche future su tali argomenti potranno avere importanti ricadute in entrambi gli ambiti industriali.

and considering the existence of dedicated OCR functions in common artificial vision systems.

Vice-versa, the problem of identifying the feature that constitutes the character on products marked by means of different technologies is of increasing interest and requires a more accurate definition. The final purpose of this article is to emphasize that the shape of printed characters is also defined by international standards, such as [19] [20] [21], which permit easy definition of algorithms for recognition, whereas controlling a marking process or developing a new technology can be dealt with by simply examining the feature generated in marking the character itself. An example of measuring the contrast of a character with respect to a background (with many noises) can be given in the ratio between the character signal and that of the background based on the geometric extension of the feature in the framed area and the intensity level of the signal. This example, with many others, can be mathematically defined and utilized as a parameter evaluating the performance of character-recognition algorithms in the industrial context. Where a monochromatic image is concerned one can, for example, consider the ratio between the average intensity of the pixels belonging to the character and the average for the background. The greater the variability or noise present in the background the lesser the contrast and thus the legibility. The two aforementioned problems are highly integrated since the performance of an OCR algorithm for any process stage (schematized in figure 1, part one of this article) depends on the marking technology and hence on the type of image acquired.

Conclusions

We have given an overview of the main research directions taken by developers of automatic marking and inspection systems.

*The optical character recognition (OCR) problem is widely discussed in computer research but from the industrial viewpoint the problems still need a systematic framework in order to help create commercial software able to reliably handle the preprocessing stage for a wide range of marking technologies for materials different from paper. The last part of this article introduced the concept of feature contrast. In product marking the contrast problem requires appropriate treatment both for the purpose of **controlling the quality** of the marking done and for defining requisites for an **automatic system**. Future research into these subjects could have important fallout in both industrial spheres.*

Notes

* Michele Lanzetta^{1,2}, Federica Fanti¹, Giovanni Tantussi¹:

1. University of Pisa, Department of Mechanical, Nuclear and Production Engineering (DIMNP), Production Section, Via Bonanno Pisano 25/B, 56126 Pisa, Italy, lanzetta@unipi.it, tel. 050-9130.22;
2. National Research Council (CNR), A. Faedo Information Technologies and Sciences Institute (ISTI), Via G. Moruzzi 1, 56124 Pisa, Italy.

Studi & Ricerche

STUDIES & RESEARCH

Note

* Michele Lanzetta^{1, 2}, Federica Fanti¹, Giovanni Tantussi¹:

1. Università di Pisa, Dipartimento di Ingegneria Meccanica, Nucleare e della Produzione (DIMNP), Sezione Produzione, via Bonanno Pisano 25/B, 56126 Pisa, Italia, lanzetta@unipi.it, tel. 050-9130.22;
2. Consiglio Nazionale delle Ricerche (CNR), Istituto di Scienze e Tecnologie dell'Informazione (ISTI) "A. Faedo", via G. Moruzzi 1, 56124 Pisa, Italia.

Ringraziamenti

Il presente lavoro è stato sviluppato come progetto didattico di Federica Fanti nell'ambito del corso di Automazione dei Processi Produttivi del corso di laurea specialistica in Ingegneria dell'Automazione, anno accademico 2006-07, presso la facoltà di Ingegneria dell'Università di Pisa. Michele Lanzetta e Giovanni Tantussi sono docenti del settore scientifico disciplinare Tecnologie e Sistemi di Lavorazione presso la medesima Università. Il progetto è stato co-finanziato dal Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) e dall'Università di Pisa come Progetto di Ricerca Scientifica di Rilevante Interesse Nazionale (PRIN, 2004).

Acknowledgements

This study was developed as an educational project by Federica Fanti in the context of a course on automating production processes for the degree course in Automation Engineering, academic year 2006-07, Engineering Faculty, University of Pisa. Michele Lanzetta and Giovanni Tantussi are professors in the scientific sector of Production Technologies and Systems at the same university. The project was co-financed by the Ministry of Education, Universities and Research (MIUR) and by the University of Pisa as a Scientific Research Project of National Importance (PRIN, 2004).

Bibliografia / Bibliography

- [14] Y. Lu, *Machine Printed Character Segmentation – An Overview*, Pattern Recognition 28 (1): 67-80 Jan. 1995.
- [15] H. Fujisawa, Y. Nakano, K. Kurino, *Segmentation Methods for Character-Recognition – from Segmentation to Document Structure-Analysis*, Proceedings of the IEEE 80 (7): 1079-1092 Jul. 1992.
- [16] Y. Zhong, K. Karu, A.K. Jain, *Locating Text in Complex Color Images*, Pattern Recognition 28 (10): 1523-1535 Oct. 1995.
- [17] R.G. Casey, E. Lecolinet, *A survey of methods and strategies in character segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18 (7): 690-706 Jul., 1996.
- [18] O.D. Trier, A.K. Jain, T. Taxt, *Feature extraction methods for character recognition – A survey*, Pattern Recognition 29 (4): 641-662, Apr. 1996.
- [19] ISO 1073-1:1976. Ed. 1, Alphanumeric character sets for optical recognition – Part 1: Character set OCR-A – Shapes and dimensions of the printed image.
- [20] ISO 1073-2:1976. Ed. 1, Alphanumeric character sets for optical recognition – Part 2: Character set OCR-B – Shapes and dimensions of the printed image.
- [21] ISO 1004:1995. Information processing – Magnetic ink character recognition – Print specifications.